DOCUMENT RESUME

ED 072 689 FL 003 734

AUTHOR	Lotz, John
TITLE	Script, Grammar, and the Hungarian Writing System.
INSTITUTION	Center for Applied Linguistics, Washington, D.C.; Hungarian Academy of Sciences, Budapest. Linguistics Inst.
SPONS AGENCY	Ford Foundation, New York, N.Y.; Hungarian Academy of Sciences, Budapest.
PUB DATE	72
NOTE	43p.; Hungarian-English Contrastive Linguistics Project Working Papers 2
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Alphabets; *Applied Linguistics; Comparative Analysis; Contrastive Linguistics; English; Grammar; Graphemes; Handwriting; *Hungarian; *Language Research; Linguistic Theory; Morphemes; Morphophonemics; *Phonemic Alphabets; Phonetic Transcription; Phonology; Punctuation; Second Language Learning; Semantics; *Written Language

ABSTRACT

This booklet forms a part of the Hungarian-English Contrastive Linguistics Project which is concerned with investigating the differences and similarities between these two languages with implications for second language acquisition. The papers here deal with the Hungarian writing system. Initial remarks concern the relationship between script and grammatical theory and how the consideration of script is significant for a complete description of a natural language; the author includes a select bibliography on script and language. The second part considers Hungarian script in terms of its basic elements or graphemes. The author outlines a systematic treatment for the graphematic study of a language considering grapheme inventory, signal formation in words and texts, and pragmatic aspects of script. The method for converting the native orthography of a language, in this case Hungarian, into a phonemic transcription is demonstrated and a sample text is provided. The concluding paper deals with the Hungarian imperative in both its spoken and written form. (VM)

ED 072689

THE HUNGARIAN-ENGLISH
CONTRASTIVE LINGUISTICS PROJECT

WORKING PAPERS

2

JOHN LOTZ

Script, Grammar and the Hungarian Writing System

FL 003 734

LINGUISTICS INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES
AND
CENTER FOR APPLIED LINGUISTICS

1972

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE HUNGARIAN-ENGLISH CONTRASTIVE LINGUISTICS PROJECT

Working Papers

EDITORS

László Dezső William Nemser

CONSULTING EDITORS

Lajos Tamás John Lotz

The Hungarian-English Contrastive Linguistics Project is jointly administered by the Linguistics Institute of the Hungarian Academy of Sciences and the Center for Applied Linguistics of Washington, D. C. The Project is jointly supported by the Ford Foundation and the Hungarian Academy of Sciences.

The major research objective of the Project is the systematic large-scale investigation of differences and similarities between the Hungarian and English languages with implications for the acquisition of English by Hungarians and the acquisition of Hungarian by speakers of English.

The Project publication, *Working Papers*, makes available research results, theoretical studies, progress reports, sample pedagogical materials and other materials relevant to Project objectives.

Communications should be addressed to Dr. László Dezső, The Hungarian-English Contrastive Linguistics Project, Linguistics Institute, Hungarian Academy of Sciences, Budapest V., Szalay utca 10—14, Hungary.

Table of contents

PART I

1. Script in the Framework of Grammatical Theory 7
2. Select Bibliography on Script and Language 13

PART II

3. Hungarian Script 17
4. The Conversion of Script to Speech as Applied to Hungarian 27
5. The Imperative in Hungarian 41
(Spoken and Written)

Bibliographical Note 48

PART I

4/5

SCRIPT IN THE FRAMEWORK OF GRAMMATICAL THEORY

In describing a language, or all the languages of the world, descriptions in modern linguistic practice aim primarily at the spoken language. If, however, we intend to cover the entire scope of natural languages, we should include other symbolic systems as well. By natural language we mean a symbolic system which can generate topically unlimited discourse. Its characteristics are: a large set of basic signs (morphemes) which can be changed and expanded; the formation of sentences by intricate hierarchic syntactic rules out of these morphemes; the formation of discourse by a loose repetition of such sentence structures; and nonlimited semantic coverage correlated with this formal arrangement. Such symbolic systems can produce an unlimited number of discourse events, can cover any known situation, and are capable of being restructured. Natural languages, so defined, may use visual rather than oral transmission as the characteristic means of communication, such as in sign languages. And, more importantly, in many language systems the spoken language co-occurs with written language. For these languages script is a normal and natural means of communication.

In societies where script is used, it is a significant medium of expression. It is the only source of language information for older stages of languages and for extinct languages where the phonological aspect has to be inferred and is less well understood. But even in modern societies, script represents major social, political and cultural forces (the choice of the kind of alphabet, e.g. Cyrillic for minor languages in the Soviet Union and Latin script in other parts of the world; the change from one type of script to another, e.g. introduction of Latin script for Turkish in the 1920's; the use of different kinds of script, such as Latin and Cyrillic for Serbo-Croatian, etc.). For languages where script is used, the spoken language alone cannot be regarded as the full representation of the language, e.g. spoken Japanese is hardly a sufficient and self-contained modality of Japanese.

Linguistic theories on the whole have not dealt adequately with this problem of script. In American structuralist tradition, script was „out-defined" by postulating that writing is not language (this results in the consequence that the journal of the Linguistic Society of America, *Language*, is not in English). In various European structuralist traditions script was touched upon unsystematically and mostly with reference to stylistic characteristics of written language. Glossematics did allow the linguistic form to appear in various substances, including speech and script; these modalities, however, were never investigated, nor was their relationship to each other clarified. In many traditional grammars a chapter on script was loosely inserted, but without integrating it in the total grammatical framework. Books which deal with script in general, such as the works of Jensen or Gelb, put the emphasis on script itself, and its relationship to the total context of language was less of a concern.

— —

For my own attempt to integrate script into the total framework of language description see Part II of this volume which treats script as a self-contained system by establishing in detail the total inventory of graphic elements in Hungarian script and the principles involved in text

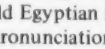
formation by means of these elements. This endeavour was based on the following general assumptions:

1) Script should be treated on its own terms and cannot be derived entirely from speech, since script contains elements which are not present in speech, such as capitals, abbreviations, hyphenated forms in modern languages and determiners in Old Egyptian hieroglyphs. It is equally true that certain features of speech are not represented in script and, therefore, script is not sufficient to substitute for a phonological description either. It is interesting to note that script seems to be closer to „underlying" morphological forms than a phonological representation.

2) In the total framework of language script has to be established in three relationships: the relationship of script to semantic coverage; the relationship of script to morphemic units; and the relationship of script to speech.

The units in script with relation to sound can either be references to single phonological elements, such as Finnish *m* corresponds to the sound /m/, or to complex phonological segments, such as Greek ψ corresponds to the cluster *ps*, or to syllable-like segments, such as Japanese 不 corresponds to *fu*. Sometimes graphic elements correspond to prosodic features, such as in Hungarian the accent on *á* corresponds to the prosodic feature of length. One can even go further and claim that the umlaut-sign on the German vowels *ä, ö, ü* corresponds to a raised second formant contrasting them with back vowels.

Units of script referring to morphemes appear wholesale in logographic scripts, such as in Chinese. In our script system the numbers or the symbol & do not correspond to any sound; they refer globally to the morpheme.

Elements of script which refer to meaning directly occur in the determiners which are common in Old Egyptian hieroglyphs or in Sumerian cuneiform, e.g. in  the upper sign refers to the pronunciation /p.r/ and the lower part , the legs, to the semantic feature of going.

Sometimes the relationship is complex, e.g. the interrogation sign ? in Hungarian indicates that the sentence is not assertive, but questioning. It can be regarded as a global morpheme, and it also corresponds to certain intonation patterns in speech.

3) Though script and speech have to be treated as separate systems for adequate description, there are certain features in the relationship between the two which are not simply juxtapositional.

It is customary to regard speech as the only normal medium of language and view script as secondary. This approach is especially prevalent in American structural lingustics, but such views occur already in Aristotle's thinking about language in his formulation which establishes the chain: psychological content → speech → script.

> *Ἔστι μὲν οὖν τὰ ἐν τῇ φωνῇ τῶν ἐν τῇ ψυχῇ παθημάτων σύμβολα, καὶ τὰ γραφόμενα τῶν ἐν τῇ φωνῇ.*

(*De Interpretatione,* Loeb; p. 114. Note the term *grapheme.*)

The arguments advanced to motivate such a position are of two kinds: a) of genetic nature and b) of generality.

a) The genetic arguments can be *ontogenetic,* i.e. the individual in his life span learns to speak first and writing is acquired later; *phylogenetic,* i.e. in the history of any social group, or mankind in general, speech developed first and script came later; or *psychogenetic,* i.e. it is claimed that in each reading or writing event there is a mediation through the underlying spoken language.

b) The arguments with reference to generality are of three kinds: *literacy,* i.e. there are no communities without a spoken language, whereas there are many communities which have no writing system; *educational,* i.e. even in communities where writing exists, not all adults are literate; and *generational,* i.e. children speak before they learn to write.

It seems clear that the genetic arguments are inconsistent with the general procedure which is followed in a structural analysis of speech. Also, the arguments of generality, valid as they may be, do not give any systematic and concrete analysis of script as an independent phenomenon. There are elements in the writing system, such as capitals, which have no correspondence in the spoken language and vice versa. Therefore, we follow the procedure to describe both media by themselves and then state the correlation between them.

It must be noted, however, that speech has basic *inherent* characteristics which give it the pivotal unmarked position among the different media of language communication, including script. The *unidirectedness* of language communication is a constitutive characteristic feature of speech, present both in articulation (where it is characterized by a dynamic manipulation of the enclosed air mass in the three-dimensional vocal tract) and in the acoustic sphere (where it is characterized by a limited pressure variation within a single dimension of oscillation dispersing in space), both of which are organized along a single time dimension. In script, on the other hand, there is no internal reason to choose this restricted linear way of communicating, since a surface can be utilized more variedly than just in linear sequences as, for example, in pictographic writings which were forerunners of the hieroglyphs in Egyptian. Moreover in speech the distribution of sounds along the syntagmatic axis, the syllabic vowel-consonant interplay, has its internal motivation, i.e. the opening-closing of the vocal tract; no such motivation exists for the distribution of letters in writing. Also, speech is a brief restructuring of the atmosphere by a low energy vibratory impact of the vocal tract, whereas writing presupposes a material background (a piece of paper, a slab of stone) and requires additional tools and materials (ink, stylus, pen). Primacy of speech as to its relationship to script cannot be viewed as evolutionary or social in the structural sense; it is „logical" because of its material inherency.

In addition, speech has certain advantages in the efficiency and immediacy of communication. It is produced with the human body alone, without any tool; it is independent of light, and can be used day and night; it fills the entire space around the source, the speaker, and does not require a straight line of connection with the receiver; it can also be greatly varied in energy from intimate whisper to long distance shouting; and it involves a very small amount of energy leaving the body of the speaker almost entirely free for other simultaneous activities.

The time sequencing of speech is reflected in the spatial-linear sequencing of script. In actual fact, however, the sequence is not strictly linear in either case. In script the graphematic elements follow after each other in thin narrow lines in various directions (Latin right to left horizontally, Arabic left to right horizontally, Chinese vertically in left to right columns) allowing limited co-occurrence of elements (Middle High German $\ddot{a} - \ddot{a}$). In speech a small number of discrete distinctive features co-occur with constructive markers (e.g. intonation patterns). So in reality the linear sequence in speech and script has to be interpreted as unlimited expandability in one dimension of symbolic components within the circumscribed range of a narrow „ribbon".[1]

The following two figures sum up what has been said about script: Figure 1 places script in the framework of the total design of language; Figure 2 lists the different types of graphemic references.

1. Cf. my article „Speech and Language" in *The Journal of the Acoustical Society of America*, Vol. 22, No. 6, 1950, pp. 712–717.

SYMBOLS

Γ graphic system
ϕ phonological system
μ morphemic (formal) sign
σ semantic reference
⌒ the basic constitutive relation in language
- - - relation from morphemic units to expressive media
⟶ graphic elements in relation to other elements

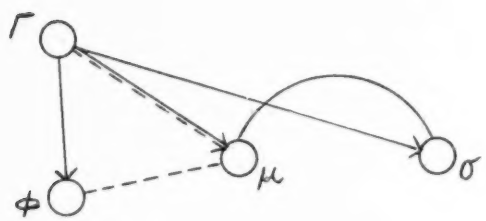


Figure 1. Language Model with Built-in Script Component.

Graphic Symbol	Phonological Reference	Morphemic Reference	Semantic Reference	Phonological Correspondence
Finnish m	/m/	–	–	/m/
Greek ψ	/ps/	–	–	/ps/
Japanese フ	/fu/	–	–	/fu/
English &	–	and	–	/and/
Egyptian hieroglyph	/f/ / /r/	n.f.r	–	/nefer/
Egyptian hieroglyph	–	p.r	–	/per/
Egyptian hieroglyph	–	p.r	'Category of going'	/per/
Hungarian ?	/Different intonation patterns/	Sentence qualifier	'Question'	/Different intonation patterns/

Figure 2. Chart of Graphic References.

SELECT BIBLIOGRAPHY ON SCRIPT AND LANGUAGE

The literature about script is extensive and varied in the extreme, both as to topical approach and soundness in scholarship. The subject has been dealt with by representatives of a number of fields: archeologists interested in ancient objects, ethnographers interested in culture, printers and artists interested in print and calligraphy, educators and missionaries interested in developing writing systems and literacy, and, of course, philologists and linguists.

The linguistic approach – which is our main concern here – has been mostly philological, mainly dealing with historical questions, especially the origin and diffusion of script. In grammars script was generally treated as a topic outside of the central core of language research. For the most part Aristotle's view that script is a secondary reflection of speech which in its turn directly mirrors meaning has been the prevalent one, as, for instance, in all schools of American structural linguistics. In the last item of this Bibliography I tried to develop a theory which integrated script organically into the general framework of grammar.

The chronological bibliography which follows represents what I consider the most significant contributions in Western scholarship dealing with the problem of script. It includes both general and comprehensive treatments of script as an independent phenomenon as well as a few articles which reflect the views of structural linguists on the role of script.

REFERENCES

von KLAPROTH, H.J. *Aperçu de l'origine des diverses Écritures de l'ancien Monde.* Paris, 1832.

FAULMANN, K. *Illustrirte Geschichte der Schrift.* Wien–Pest–Leipzig: A Hartleben's Verlag, 1880.

TAYLOR, I. *The Alphabet.* 2 volumes, 2nd edition, London: Kegan Paul, French and Co., 1899. (1st edition, 1883).

WEULE, K. *Vom Kerbstock zum Alphabet.* Stuttgart, Kosmos: Gesellschaft der Naturfreunde, 1915.

JENSEN, H. *Die Schrift in Vergangenheit und Gegenwart.* Revised edition, Berlin, VEB Deutscher Verlag der Wissenschaften, 1958. (1st edition, Glückstadt um Hamburg: Verlag von J. J. Augustin, 1935; 1st version, *Geschichte der Schrift.* Hannover, 1925.)

PEDERSEN, H. *Sprogvidenskaben i det Nittende Aarhundrede; Metoder og Resultater.* Copenhagen: Gyldendalske Boghandel, 1924. (Vol. 24 in the series *Det Nittende Aarhundrede;* translation *Linguistic Science in the 19th Century.* Cambridge: Harvard University Press, 1931; republished in paperback under the title *The Discovery of Language.* Bloomington: Indiana University Press, 1962.)

LEJEUNE, M. „La language et l'écriture" in *L'évolution humaine* volume III Paris, 1934.

FÉVRIER, J.G. *Histoire de l'écriture.* 2nd edition. Paris: Payot, 1959. (1st edition, 1948).

GELB, I.J. *A Study of Writing – Foundations of Grammatology.* 2nd edition. Chicago: University of Chicago Press, 1958. (1st edition, 1952.)

OGG, O. *The 26 Letters.* New York: The Thomas Y. Crowell Co., 1961.

DIRINGER, D. *The Alphabet – A Key to the History of Mankind.* 2 volumes, 3rd edition (revised with the assistance of Reinhold Regensburger). New York: Funk and Wagnalls, a Division of Readers Digest Books,

Inc., 1968. (1st English edition, Hutchinson and Company Ltd., 1948; 1st version *L'alfabeto nella storia della civilta.* Firenze, 1937.)

ULDALL, H.J. „Speech and Writing," in *Acta Linguistica,* vol. IV, 1944, pp. 11–16.

VACHEK, J. „Some Remarks on Writing and Phonetic Transcription" in *Acta Linguistica,* vol. V, 1945–49, pp. 86–93.

LOTZ, J. „The Role of Script in Describing the World's Languages," in *Visible Language,* vol. V, number 1, 1971, pp. 75–81. (Excerpted from a presentation at the Burg Wartenstein Symposium „Toward the Description of the Languages of the World" held August 1–8, 1970).

This bibliography was compiled for a conference on reading problems sponsored by the National Institute of Child Care and Human Development in May 1971 at Belmont, Maryland.

PART II

HUNGARIAN SCRIPT

Hungarian, like languages on its cultural level, occurs normally in two media: in *speech* and in *script*. Other varieties of communication in Hungarian, such as the tactile communication of the deaf mutes, and the transposition of one medium into another, as for instance the transposition of normal writing into Morse code, is of no interest to us in this connection.

– . –

Systematic treatment of script involves: the inventory of the elements called „graphemes"; and the formation of texts with these graphemes both with reference to a text and the mechanical formation of this text as to pages., lines, etc. Graphemes are glyphic or punctual elements. They are definable with reference to a surface (called reference frame) as to position and shape. The reference frames are usually homogeneous (but cf. italics or different sizes as in m^2). In handwritten form great variations can occur in script.

The study of writing will be called *graphematics*. (The word *graphology* which could be a technical term in analogy with *phonology* is commonly used for personality analysis based on handwriting and therefore not usable.)

The most effective way of accounting for the inventory of Hungarian graphemes is to list them according to functions. As in phonology, where the elements are auditory, there are two kinds of elements in Hungarian script: a) those which function as constituents of the morphemes in texts; and b) those which indicate the structuring of the texts, the interpunctuation signs. The morphographemic elements are of two kinds: in analogy with the spoken language, we might distinguish units which constitute morphemes themselves (global signs), and elements which serve as components in meaningful units (letters). These differ from similar categories in speech, since in script global-marginal elements are abstract in reference, whereas in speech they are always interjectional.

Thus, a systematic treatment of graphematics would follow the plan:

HUNGARIAN SCRIPT

Introduction

I. Structure of Script
   A. Grapheme Inventory
      1. Signal Constituents
         a. Letters
         b. Global Signs
      2. Punctuation Marks
   B. Signal Formation
      1. The Graphic Word
         a. Simple Words, Compounds, Hyphenated Complexes
         b. Use of Capital Letters
         c. Abbreviations
         d. Mechanical Cuts (hyphenation)
      2. Text Formation
         a. Sentence Formation
         b. Higher Units (paragraph, chapter, etc.)
         c. Mechanical Arrangements (page, etc.)

II. Pragmatics of Script
   A. Handwriting and Typing/Printing
   B. Emotional Features (italics, capitalization)
   C. Style of Letters

GRAPHEME INVENTORY

Signal Constituents

Letters

Basic Shapes

The letters are the basic glyphic shapes; they have no definite meaning in themselves; they function as constitutive elements, out of which, either directly, or through the mediation of certain letter combinations (digraphs, trigraphs), words are constructed. (They may carry meaning only in abbreviations as substitutes for full words, or when the meaning refers to the shape of the letter, e.g., *X-lábú* 'knock-kneed', *L-alakú* 'L-shaped'.) They occur in two corresponding shapes: small letters and capitals (or, in typewriter usage, lower and upper case), both in handwriting and in printing. It would be possible to regard capitalness as a special feature added to small letters. The glyphic shapes are arranged traditionally (Roman alphabet) in a sequence:

a b c d e f g h i j k l m n o p q r
A B C D E F G H I J K L M N O P Q R

s t u v w x y z
S T U V W X Y Z

(Additional shapes, such as Danish *æ* and *ø* are hardly ever used in Hungarian texts; usually suitable substitutions are made, associated with letter shapes and sound.)

Diacritic Signs and Vowel Letters

The following diacritic signs occur in Hungarian above a middle-line letter:

(·) ´ ¨ ˝

(·) is put in parenthesis because it occurs as part of a small *i* and *j* without distinctive value, never on capitals. Note, however, that *Í* with acute accent does occur, e.g. *Írország* 'Ireland'. (Therefore in English where the raised dot is the only diacritic sign, it is counted as an integral part of a letter and there is no need for a special section on diacritics.)

The diacritic signs, disregarding · (raised dot), are added to the following set of letters:

a e i o u

resulting in the following set, to be called *vocalic* letters:

a, á, e, é, i, í, o, ó, ö, ő, u, ú, ü, ű

(This chance accentuation allows a formal definition of vowel letters in Hungarian without recourse to substantive considerations.)

The remaining letters will be called *consonant* letters.

On capital *I*, both printed and written, the accent is often omitted, though its use is required by orthographic rules.

The letters *á, é, í, ó, ő, ú, ű,* will be called *accented,* the others *unaccented.* (The diacritic acute signs correspond to length in phonology.)

In foreign words used in Hungarian texts, other combinations and other diacritic marks may occur: *ä,* or *å* (in Swedish), Danish *ø*; and also consonants may be combined with diacritic marks: e.g. French *ç*, Czech *č*, Spanish *ñ*, etc., but these cannot be regarded as part of the Hungarian graphemic inventory.

The Alphabet

The units of script in Hungarian traditionally include, besides the single graphic shapes listed above *(hengraph)*, and the vowel letters with their appropriate diacritic signs, a number of letter combinations. These letter combinations consist either of two letters, called *digraphs*, or – in a single case *dzs* – of three letters, called *trigraph.* The digraphs and trigraph function as units in Hungarian orthography, not only through their monophonematic correspondence in the speech domain and in their distributional similarity to single letters, but also from a strictly graphemic point of view as they are handled as units like the single letter graphemes in hyphenation: e.g. me-se ‚fairy tale', me-sze ‚his whitewash'.

The graphemes, now including the polygraphs and the accented vowels as well, are listed in a traditional Hungarian sequence called alphabet. Each letter can be referred to by its name. This sets up for each grapheme a morpheme as well. (The graphemes, however, do not primarily refer to phonemes, as often assumed.)

THE HUNGARIAN ALPHABET

(commonly occurring foreign letters are in parenthesis)

Letter	Name of letter in Hungarian Orthography	Name of letter in phonemic transcription
a	a	ɔ (ɔ̄)[5]
á	á	ā
b	bé	bē
c	cé	cē
cs	csé	c̆ē
d	dé	dē
dz[10]	dzé	ʒ̄ē
dzs[10]	dzsé	ʒ̆̄ē
e	e	ɛ (ä)[2] (ɛ̄) (ǟ)[5]
é	é	ē
f	eff	ɛff (äff)
g	gé	gē
gy	gyé	ɟ̄ē
h	há	hā
i	i, rövid í	i, rövid ī
í	í, hosszú i	ī, hossu ī
j	jé	jē
k	ká	kā
l	ell	ɛll (äll)
ly	ell-ipszilon(elly)	ɛll-(äll-)ipsilon, (ɛjj, äjj)[3]
m	emm	ɛmm (ämm)
n	enn	ɛnn (änn)
ny	enny	ɛn̄n̄ (än̄n̄)
o	o, rövid ó	o, rovid ō
ó	o, hosszu ó	ō, hossu ō
ö	ö, rövid ő	ö, rövid ō
ő	ő, hosszú ő	ō, hossu ō
p	pé	pē
(q)	(kú)	(kū)[7]
r	err	ɛrr (ärr)
s	ess	ɛs̆s̆ (äs̆s̆)
sz	essz	ɛss (äss)
t	té	tē
u	u, rövid ú	u, rövid ū
ú	ú, hosszú ú	ū, hossu ū
ü	ü, rövid ű	ü, rövid ū
ű	ű, hosszú ű	ū, hossu ū
v	vé	vē
(w)	dupla vé	(duplɔ vē)[7]
(x)	(iksz)	(iks)[7]
(y)	(ipszilon)	(ipsilon)[7]
z	zé	zē
zs	zsé	z̆ē

In Hungarian texts before 1908 the digraph *cz* was used instead of *c;* this is still common in names, e.g. *Czuczor.* In older texts the digraph *ds* was used instead of the current trigraph *dzs.*

In foreign words – which, however, are in common use – the digraph *ch* occurs, e.g. *jacht* 'yacht', *technika* 'technology', *mechanika* 'mechanics'.

In names further digraphs occur traditionally which play a role in hyphenation:

vocalic:	*ij*				
	aa	*aá*			
		eé			
	oo	*oó*			
	ua	*uo*			
	ee	*eő*	*eo*	*ew*	
	uu				
consonantal:	*cz*	*kh*		*dgy*	*ff*
	ch	*gh*			*ll*
	chs	*th*	*tth*		*ss,* etc.
	ts	*rh*			

Remarks

1. Small letters and capitals are indicated by the addition of the attribute *kis* 'little', and *nagy* 'big', e.g. *kis bé /kiš'bē/* 'small b', *nagy bé /nɑd'bē/* 'capital b'. The general names for the alphabet in lower and upper case are *kis ábécé* 'lower case (small) alphabet' and *nagy ábécé* 'big alphabet', respectively. In digraphs and trigraphs, except in capitalization for emphatic purposes, only the first letter is, of course, capitalized, therefore they are not listed specifically.

2. The „open" (*á*) is used for the names of the letters beginning in *e* in the pronunciation of those Hungarians who have two „e" phonemes; this goes for the name of the letter *e* as well.

3. For *ly* the use varies either (*elj*), or (*ellipsilon*); the latter seems to be more common.

4. Where the vowel qualities in pronunciation are very similar, reference to the names of accented and unaccented *i, í, o, ó, ö, ő, u, ú,* and *ü, ű,* is made by the addition of the attributes *rövid* 'short', and *hosszú* 'long', e.g. *rövid ó* 'short (unaccented) o', and *hosszú ó* 'long (accented) ó'; in pronunciation the vowel is always long. (This is not the case with *a, á,* and *e, é* where the qualities differ.) In the continuous enumeration of the alphabet, however, the attributes are missing: */o/, /ö/,* etc. In alphabetical listings (e.g. in railroad stations in timetables) *a, á, e, é* are kept separate, whereas, *i~í, o~ó, ö~ő, u~ú, ü~ű* are intermingled. (This shows native awareness of phonological qualities.)

5. In the enumeration of the alphabet, and also, when referring to them, the names of *a* and *e* are usually lengthened: /ɔ̄/ and /ɛ̄/ or (/ā/), though these sounds otherwise don't occur long in the standard spoken language (they do occur in dialects).

6. The foreign letters *q, w, x, y,* which are listed above, may occur normally in Hungarian texts. The names are borrowings and have no special designation in Hungarian, except the name of *w, dupla vé,* meaning 'double v'. Note that *y* though it occurs commonly as second element of a digraph (*gy, ly, ny, ty*), does not occur in Hungarian words independently, except that at the end of names it occurs frequently as a remnant of traditional writing, with the sound value *i.*

7. Additional foreign letters from script using Latin characters appear immediately following the shape of the similar Hungarian letter, including digraphs. Other alphabetical script systems, e.g. Cyrillic, have to be transliterated. Thus, German *ä* follows *á,* Estonian *õ* follows *ő,* Czech *č* follows *c,* Turkish *ş* follows *sz.* There are no rules for competing foreign letters, e.g. *ä* and *å,* or *ç* and *č.* (This problem occurs only in learned publications.)

8. The names of the letters are morphemes referring to the letters (and not to the sounds). In general, linguistic consciousness is mostly associated with script.

9. On Hungarian typewriters generally there is no key for *í, ú, ű.* However, *ő, ó, é, á* are never missing.

10. *dz* and *dzs* are rare.

11. In enumeration the basic shapes are used, e.g. 2a, 2b, not 2)a, 2)b.

Global Signs

There are a number of signs in script which have a clear morphemic-semantic reference and which are never used as a productive component of other morphemes. These are called *global* morphemes. (In speech, there are also morphemes of this kind. But whereas in speech they are always marginal, interjectional and asyntactic, in script they are abstract in reference.) We group them in two categories: the numbers and other global signs:

a. The numbers are:

0 1 2 3 4 5 6 7 8 9

All permutations of any order occur. For decimals a comma is used, e.g. 3, 14 (*három egész tizennégy század*).

b. Other global signs include the mathematical operators:

+, −, · or x, : or /
the sign of equivalence: =
signs such as: % percentage
§ paragraph (cf. below)
./. turn the page
° grade of temperature

In special texts, especially scientific texts, as in mathematics, astronomy, genetics, chemistry, etc., additional signs are used in various arrangements, e.g.

√ root
♂ male
♃ Jupiter
→ implies

Punctuation Marks

Punctuation signs are special markers referring to the structuring of a text composed of letters (and to a minor extent of global signs). They have no stated traditional order of presentation. In the following they will be grouped according to their joining of the letter sequences (proclitic, enclitic, amphiclitic or internal) and their function in a text.

Sentence markers
(Always enclitic)

Internal	quasi-internal	final
,	:	.
;		?
'		!
		...

Quasi-internal is always non-final, but it may be followed by either sentences or by word stretches.

The sign of interruption may be preceded by a comma : ,...

Word markers

These signs indicate interrupted coherence in words, either because of the end of the line (hyphenation) before the questioning particle *-e*, for delayed reference or for other reasons, e.g. *lesz-e* 'will it be', *nyelv- és irodalomtudomány* 'language and literary study', *meg-megáll* 'he stops (repeatedly)', *Svéd–magyar* 'Swedish–Hungarian'.

The short hyphen is either enclitic or internal, the long hyphen always internal.

Further text indicators

—— (proclitic) for direct quotation in novels (cf. long hyphen)
§ (proclitic) for a unit in strictly organized texts, e.g. legal texts.
( )(amphiclitic) for a text portion deviating in reference from main text.

Other signs referring to the mechanical organization of the text also occur such as " repetition sign, refers to the –"– corresponding position in the preceding line.

Special signs

They occur in special texts, e.g.

' ' single quotes for meaning (in linguistic works)

Some of the punctuation signs may be identical in shape, but different in function or position, e.g. —— (long hyphen can serve as word marker or indicator of quotation). (Homography.)

TEXT FORMATION

Normally a text consists of a combination of morpheme-constitutive signs and punctuation signs. Punctuation signs are missing in some cases, for instance, on street names, book titles, etc. Text formation involves the formation of meaningful texts and their arrangement mechanically in a sequence adjusted to writing space.

The mechanical aspects of text formation include the direction of the script sequence, in Hungarian from left to right, the use of reference squares, which are normally homogenous but occasionally variation in size occurs, for instance m^2, and spacing. The text then is adjusted in lines, normally a straight line, on pages and in complete books. On a page a special column arrangement is possible. Normally the lines go entirely across the page, but for special reasons there might be other arrangements, for instance in lists of names and in poetry.

From the point of view of content the graphic units are word and sentence, which parallel similar notions in speech, but beyond this there might be further groupings in paragraphs, chapters, etc., which are less clearly marked in speech.

The distribution of letters mirror similar facts in speech and will not be treated. Graphic words are always terminated by space or by some punctuation signs such as hyphen, or full stop. The initial element is either in upper or in lower case.

Higher semiotic units are sequences of words common to both speech and script. They are marked by punctuation signs of two kinds either marking internal organization, such as comma, or completed sentence such as question mark. The higher order graphic arrangements, such as paragraphs, chapters etc., are less regulated by norms than graphic sequences up to the sentence level.

NOTE

This presentation of Hungarian script from my *Hungarian Reference Grammar* prepared under an Office of Education contract, 1969, is a revision of a corresponding chapter in *Das Ungarische Sprachsystem,* Ungarisches Institut, Stockholm, 1939, pp. 17–25.

Additional materials referring to the subject have been collected, but they are not presented here. These include: the use of capitals, the function of punctuation, abbreviations, and various printing technical devices (underlining, italics, etc.)

THE CONVERSION OF SCRIPT TO SPEECH AS EXEMPLIFIED BY HUNGARIAN

This paper[1] demonstrates how a native orthography, as exemplified by Hungarian, can serve as a device for phonemic transcription. The method developed for this exercise in applied linguistics promises to be useful in language teaching. The main text treats the problem in a straightforward, mechanical manner and in explicit detail; background information and scholarly references appear in the notes.

Even if the emphasis in teaching a language is on the spoken language and the initial period of instruction is devoted solely to oral exercises, as in most modern approaches to language teaching, the need to fix the language material in written form arises early. A systematic phonemic transcription is preferable for such a purpose. In the case of an unwritten language or for languages which do not use an alphabetic system of writing, there is no problem, since the development of a phonemic transcription is clearly the only choice.[2] In the case of languages which have an established alphabetical writing system, but where the relationship between pronunciation and writing is highly complex and erratic, as in English, the introduction of a specific transcription is necessary. But what can be done in the case of a writing system which is nearly, but not entirely, phonemic? The use of two parallel writing systems, (1) the official orthography and (2) a phonemic transcription–as has been done–is a source of confusion. If one is to choose between the two graphic systems, the official orthography, an existing representation of the language which the student will have to learn and use anyhow, has precedence. In what follows, I will attempt to show how it is possible to convert the official orthography into a phonemic transcription, applying the procedure to Hungarian.

The theoretical position taken here is based on the acceptance of the two normal language media, speech and writing, as equivalent symbolic systems of expression.[3] The paper deals with the central core in language: the empty signalling elements, i.e. letters in script and sounds in speech. The relationship between the „constructive" features of the two expressive media—

1. This version of the paper was prepared under a contract between the U.S. Office of Education and the Center for Applied Linguistics (transferred from Columbia University) to produce a „Hungarian Reference Grammar". The conversion chart has been distributed in my classes at Columbia University since 1959.

   The materials presented here form a chapter in the section entitled „Expressive Media", which includes: 1. Speech (Phonology); 2. Script (Graphematics); 3. Conversion of Script to Speech (reproduced here); 4. Conversion of Speech to Script; and several chapters on a Hungarian X-Ray Sound Motion Picture Film and on various Hungarian contrastive subjects.
2. The teaching of unwritten languages may take place even in a formal classroom situation, since such languages are sometimes taught for practical purposes, e.g. in the training of Peace Corps volunteers.
3. This is the position taken in my grammar, *Das ungarische Sprachsystem*, Stockholm, 1939. A detailed analysis of the complex situation involving the imperative is carried out in my article „The Imperative in Hungarian (Spoken and Written)", see below in this volume.

   Among the various structuralist schools, glossematics took a similar position. Hjelmslev and Uldall, however, relegated the normal media, speech and script, to a lower level, called *substance*, in the glossematic hierarchy, to be determined by a higher level, called *form*. But the theory was not–and I think could not be–applied to empirical language material.

between stress, intonation and juncture in speech on the one hand, and punctuation in script on the other – is not explored. Here the interrelation is less systematic and less clearly understood than in the case of letters and sounds. (The constructive features constitute a concomitant channel in the transmission of information.) Also omitted, for obvious reasons, are features which are restricted to one of the media, such as the use of capital letters, hyphenation, and space arrangement in script, and emotional features in speech.

The paper will treat briefly (A) the graphemes[4] of Hungarian, (B) the phonemes, and (C) the relationships between the two. The point of departure is script, and it is asked how script can be converted to speech (pronunciation).[5] (The reverse problem of conversion from script to speech is not identical with that of converting speech to script.)

A. GRAPHEMES

The Hungarian writing system uses Roman letters with the addition of diacritical signs over the vocalic letters. The functional units in the Hungarian writing system, the *graphemes*, are of three types: (1) *hengraphs*,[6] consisting of a single letter; (2) *digraphs*, the combination of two letters (*cs, sz, zs, dz, gy, ly, ny, ty*); and (3) a *trigraph*, the combination of three letters (the only example is *dzs*).

The Hungarian alphabet is arranged in the following traditional order: *a, á, b, c, cs, d, (dz), (dzs), e, é, f, g, gy, h, i, í, j, k, l, ly, m, n, ny, o, ó, ö, ő, p, (q), r, s, sz, t, ty, u, ú, ü, ű, v, (w), (x), (y), z, zs. (Note: dz* and *dzs* are rare and are often disregarded in the enumeration of the alphabet. The letters *q, w,* and *x* occur only in foreign words. The letter *y* occurs in Hungarian only as the second element of a digraph [and in family names in traditional spelling].)

As mentioned above, we disregard specific features of script which have no systematic correlation in speech, such as capitalization; hyphenation;[7] abbreviations, e.g. *kb.* (for *körülbelül*) 'approximately'; logographic signs, e.g. &, %; etc.

4. I use the term grapheme rather reluctantly. But a term is needed to differentiate between the unit shapes in script, for which the term *letter* is retained in accordance with conventional usage, and the functional units in the writing system, for which the term *grapheme* is introduced. In discussions of Hungarian, the lack of this distinction leads to confusion. (I would like to add that it seems to me that the Aristotelian term referred to in Part I. refers to script in general, rather than to units in script.)
5. All citations of Hungarian material are given in italics. Hungarian materials rewritten according to the conversion rules are given in Roman. Phonemic references are enclosed within square brackets. Underlining indicates a segment to be rewritten or re-interpreted according to the appropriate conversion rule; double underlining indicates that two conversion rules are involved. In addition, the following symbols are used:

C	a consonant
C_1C_1	a geminate consonant
⇒	to be respelled or re-interpreted as ...
↱	voicing switch before an element of
↳	the opposite set, and *h*
ϕ	zero

6. In a script system like Hungarian there is the need for a term for functional units consisting of a single letter. Due to my classical prejudice, I have ruled out *unigraph* because it combines a Greek and a Latin root. (But the word „television" does not disturb me.) Of the available Greek morphemes referring to one, *mono* was obviously out because of *monograph;* the choice of the femine form, *mia,* seemed to be unmotivated; so I settled on *hengraph.* (Note that *hendecasyllabic* is a recognized term in English metrics.)
7. Hyphenation might be distinctive, e.g. *fe-lül* 'above' (single word) vs. *fel-ül* 'sit up to' (compound word).

B. PHONEMES

Hungarian has the following phonemes (no attempt is made here at a comprehensive analysis, only modes of articulation being indicated; notation is in traditional transcription):[8]

Vowels:[9] short: u o ɔ (a) i ü ö
long: ū ō ā ī ē ű ő
Semivowels:[9] j (w) h
Oral stops: p t k t' b d g d'
Nasals: m n ń
Fricatives: f s š v z ž
Affricates: c č ʒ ǯ
Liquids: l r

As mentioned above, stress, intonation, and juncture will not be included.

C. CONVERSION OF LETTERS TO SOUNDS

There is a high degree of correlation between letters and speech sounds in Hungarian. The orthography is based essentially on two principles:[10] the *phonemic* principle, which assigns a single sound value to the graphemes, including the digraphs and the trigraph; and an overriding *morphophonemic* principle, according to which at some morphemic junctures the underlying form is written. In addition, there are *idiosyncratic* elements in script.[11]

8. This transcription is the one commonly used in American linguistics. It is for all practical purposes identical with the one devised in the 1870's by Joseph Budenz, the great codifier of Uralic linguistics.
9. The short, open, unrounded [a] and the labial semivowel [w] are marginal; they occur, though, normally in such „foreign" nominal roots as *Svájc* 'Switzerland' [švajc] and *autó* 'automobile' [ɔwtō]. They are discussed for the purposes of this paper under Rule 11.

   Since the orthography does not indicate the closed [E] ↔ open [ä] differentiation this distinction will be ignored in this paper.
10. Latin script for Hungarian was introduced in the eleventh century. (Earlier, Hungarian words appeared in Arabic geographical descriptions and Byzantine political tracts. There was also a native runic script in existence, based on the Turkic runes, which survived until modern times.) The influences shaping Hungarian orthography were: Neo-Latin-Italian influence (*gy* for palatal [d'], cf. Italian *maggiore;* and the letter *s* for [š], reflecting North Italian pronunciation); Czech-Hussite influence through the mediation of the University of Cracow in Poland (indication of vowel length); and German (the „Umlaut" diacritic, e.g. *ü* and *sz* for [s]). The orthography became fairly well stabilized in the sixteenth century through the typographic practices of the Protestant printing presses, e.g. utilization of morphophonemic writing. During the Counter-Reformation, Catholic printers used a fairly strict phonemic orthography, but around 1800 the etymological (= morphophonemic) principle prevailed. (In a general sense it can be said that the conversion rules set forth here result from a consistent application of the „Catholic" orthographic rules instead of the prevailing „Protestant" ones.) After its establishment in the 1830's, the Hungarian Academy of Sciences became responsible for the orthography. The regulations of the Academy have the force of a government decree for schools and for official use. The latest regulations were issued in 1954: *A magyar helyesírás szabályai* [Rules of Hungarian Orthography], Budapest, 1954. For a succinct account of the above, see J. Kniezsa, *A magyar helyesírás története* [History of Hungarian Orthography], 2nd rev. ed., Budapest, 1959.
11. New idiosyncrasies are sometimes created by the Academy itself. For instance, the 1954 regulations require that Russian *a* be rendered by Hungarian *a* [ɔ], except in the name „Stalin", where *á* [ā] is to be used, *Sztálin.* Also the name „Lenin" is spelled *Lenin* and not **Lenyin,* which would normally be required by the transliteration rules. Similarly the selfdesignation of the Ostyaks (a closely related people to the Hungarians in Western Siberia) is idiosyncratically written Chanti, instead of Khanti.

Our purpose is to show that segments of the text written in the standard orthography can be converted into phonemic representations by a simple set of conversion rules. In the following we shall distinguish two sets of correspondences: (I) the case of the application of the *phonemic* principle, where general conversion rules indicate how the letters are to be pronounced; and (II) specific instances of *nonphonemic* conversion rules, which indicate how segments of the text written nonphonemically are to be rewritten in terms of the phonemic principle. The second set of correspondences treats cases involving the morphophonemic principle (which is by far the most important part), as well as a few cases of phonological deviation and of idiosyncratic spelling. The major portion of any Hungarian text, as can be seen from the sample text presented below, is covered by the phonemic principle.[12]

I. The phonemic principle. There is a basic one-to-one correlation between grapheme and phoneme, as shown below. (The grapheme is given first, matched by the symbol for its transcription.)

a	[ɔ]	*f*	[f]	*m*	[m]	*sz*	[s]
á	[ɑ̄]	*g*	[g]	*n*	[n]	*t*	[t]
b	[b]	*gy*	[dʲ]	*ny*	[ń]	*ty*	[tʲ]
c	[c]	*h*	[h]	*o*	[o]	*u*	[u]
cs	[č]	*i*	[i]	*ó*	[ɔ̄]	*ú*	[ū]
d	[d]	*í*	[ī]	*ö*	[ɔ̈]	*ü*	[ü]
dz	[ʒ]	*j*	[j]	*ő*	[ɔ̈̄]	*ű*	[ṻ]
dzs	[ǯ]	*k*	[k]	*p*	[p]	*v*	[v]
e	[ɛ]	*l*	[l]	*r*	[r]	*z*	[z]
é	[ɛ̄]	*ly*	[j]	*s*	[š]	*zs*	[ž]

Note: The graphemes *j* and *ly* have the same phonemic value. Consonant gemination is indicated by repetition of the letter in the case of hengraphs, and by repetition of only the first letter in the case of digraphs and the trigraph, except in hyphenation and compounding, when the entire grapheme is repeated, e.g. *hattyu* 'swan' [hɔtʲtʲu], hyphenated *haty-tyu;* and *jegygyürű* 'engagement ring' [from *jegy* 'sign (engagement)' + *gyürű* 'ring']; the graphemes *dz* and *dzs* are shown orthographically as geminates only in repetitive suffixation, e.g. *eddzük* 'let us train' (cf. *edz* 'train').

The graphemes on the basis of their phonemic correspondence are divided into two categories:

1. Vocalic graphemes (vowels):[13] a á e é i í o ó ö ő u ú ü ű
2. Consonant graphemes (consonants): all the other graphemes.

Except for the marginal case under Rule 11 below, only consonants are affected by the following non-phonemic conversion rules.

II. Non-phonemic rules. Deviations from the phonemic interpretation of the graphemes are of three kinds: (a) morphophonemic, (b) phonological, and (c) idiosyncratic.

12. The use of the phonemic principle in Hungarian orthography applies to segments which are *unmarked,* to use a term which has recently become fashionable in America.

    The relationship between grapheme and sound is bi-unique except in the following three cases: the graphemes *j* and *ly* both indicate [j]; the grapheme *a* corresponds to [ɑ̄] and [a]; and the grapheme *u* corresponds to [ū] and [w] (cf. conversion rules 8 and 11).

13. Distributionally inclined formalists among structural linguists would use for the definition of the class of vocalic letters in Hungarian the fact that these are the letters capable of accentuation (the dot on *i* and *j* can be ignored), claiming thereby a superior formal definition. I would regard such a definition as an empty „trick"; the classes would not change, even if they were set up by enumeration. (Such a formal definition would not work for Finnish, where *ö* represents [ö] and *y* represents [ü].) The basis for the distinction is in speech articulation, not in the nature of the script.

A. MORPHOPHONEMIC ALTERNATIONS

1. Degemination. Preceded or followed by a consonant, all geminates are simplified. Geminates may occur either as the first two letters or the last two of a multiconsonantal (normally three-consonantal) cluster,[14] e.g. *szebbnél* → szebnél 'at the more beautiful', *házunkkal* → házunkal 'with our house', *rossztól* → rosztol 'from the bad'.

Symbolized: $C_1 C_1 \rightarrow C_1$ when preceded or followed by another consonant.

2. Voicing switch. Obstruents participating in the voiced-voiceless correlation are changed to conform with the voicing of the initial obstruent of the following morpheme, i.e. before stops, affricates and fricatives (but not before nasals, liquids or vowels);[15] e.g.:

rész [rēs] 'part'	*réz* [rēz] 'copper'
résztől [rēstől] 'from part'	*réztől* [rēstől] 'from copper'
részben [rēzbɛn] 'in part'	*rézben* [rēzbɛn] 'in copper'

(But cf. *résznél* [rēsnēl] 'at the part', *réznél* [rēznēl] 'at copper'.) Before *h* devoicing takes place, e.g. *adhat* 'he may give' (*d* → t).

The switching symbol indicates that the grapheme is to be replaced by its matching partner.

Symbolized:

p t k ty f sz s c cs	↵	before a grapheme in the
b d g gy v z zs dz dzs	↳	switching set and *h*

3. Adpalatalization. The graphemes *t, d, n* and *ty, gy, ny* followed by a *j* are pronounced as short palatals when preceded by a consonnant, otherwise as geminate palatals, e.g. *látja* → láttya [lăt′t′ɔ] 'he sees it'; *hagyja* → haggya [hɔd′d′ɔ] 'he leaves it'; *bántja* → bántya [bănt′ɔ] 'he hurts it'.

Symbolized:

t	*ty*		ty after a consonant, tty otherwise
d	*gy*	+ *j* →	gy after a consonant, ggy otherwise
n	*ny*		ny after a consonant, nny otherwise

4. Adaffrication. The graphemes *t* and *d* followed by *sz, s, z, zs,* and *c* become short affricates when preceded or followed by another consonant, otherwise geminate affricates,[16]

14. A morphophonemic sequence of three identical consonants is written as a geminate, e.g. tollal (for **toll-lal*) 'with the pen', *szebben* (for **szebb-ben*) 'in the more beautiful' (cf. *szebb-en* 'in a more beautiful way').

15. Before [h] only voiceless obstruents can occur, e.g. *adhat* → athat 'he may give'. [v] does not affect the preceding obstruent, but it is affected by the voicelessness of the following obstruent, e.g. *adva* 'given', *hatva* 'effected'; but *év* 'year', *évtől* 'from the year' → éftől. (Historically, [v] came from Old Hungarian [w] and was not part of the obstruent system.)

It is interesting to contrast English and Hungarian obstruent clusters. Superficially, they appear to be very similar, namely, they are either voiceless or voiced in their entirety, apart from a few exceptions such as *width*. The genesis of these clusters, however, is very different; cf. my „Contrastive Study of the Morphophonemics of Obstruent Clusters in English and Hungarian" in *Miscellanea di studi dedicati a Emerico Várady*, Modena, 1966, pp. 197–201.

16. It would be possible, in view of the fact that Rules 1 and 2 are already established, to state the rule in a simpler fashion only as *t* + *sz* → cc and *t* + *s* → ccs. Rule 2 would take care of the devoicing and Rule 1 of the degemination, when applicable, e.g. *bolondság* 'silliness' [bolonc̆ăg] would operate in the following way: the *d* would become *t* before *s* according to Rule 2, *ts* would become *ccs* according to Rule 4 and *ccs* would become *cs* according to Rule 1. I have not adopted this interpretation because I wanted to hold the number of operations to a minimum.

Graphic doublets can arise, e.g. *fáradság* 'endeavor' and *fáradtság* 'tiredness', from *fárad* 'to get tired, to work hard for' + *-t* (past participle) and *-ság* (abstract suffix)—both pronounced [fār ɔc̆c̆ăg] (homophony *vs.* heterography).

e.g. *ronts*[17] → roncs 'destroy', *fűtse* → fucse [fűčɛ] 'he shall heat it', *utca*[18] → ucca 'street', *szabadság* → szabaccság [sɔbɔččāg] 'freedom', *kétszer* → kéccer [kēccɛr] 'twice'.

Symbolized: $\left.\begin{matrix} t & tt \\ d & dt \end{matrix}\right\}$ + $\begin{cases} sz\ z\ c \rightarrow \text{c after or before a consonant, cc otherwise} \\ s\ zs \rightarrow \text{cs after or before a consonant, ccs otherwise} \end{cases}$

Sometimes the phenomena described above (Rules 1–4) occur across word boundaries, e.g. *egy ház* → ety ház 'one house'.

5. Laryngal loss. In root-final position *h* is not generally pronounced, except before a vowel,[19] e.g. *méh* → mé 'bee', *méhben* → mében 'in a bee'; but *méhek* 'bees'.

Symbolized: *h* → φ in root-final position, except before a vowel.

6. Pseudodigraphs. Sometimes grapheme sequences which appear to constitute digraphs, single or geminate, actually belong to separate morphemes and must be pronounced accordingly (underlining indicates the element which is a part of the separate morpheme), e.g. *község* 'village, community', from *köz* 'general' + *-ség* (abstract suffix); *egészség* 'health', from *egész* 'whole' + *-ség* (for the actual pronunciation of both *község* and *egészség*, see Rule 14, below); *meggyón* 'confess', from *meg-* (converb for completion) + *gyón* 'he is confessing'. This often occurs in the traditional spelling of proper names (cf. Rule 12, below), e.g. *Vörösmarty* → Vörösmarti (name of a poet).

Pseudodigraphs (and pseudotrigraphs) can also occur at the seams of compounds, e.g. *vadzerge* 'wild mountain goat', *lúdzsir* 'goose fat'.

Symbolized: *C* → the letter C is not part of a digraph or a trigraph.

Here, of course, respelling is not possible by means of the Hungarian orthography. The symbol → indicates re-interpretation, rather than respelling. In hyphenation, however, the distinction can be made, e.g. *köz-ség* vs. *má-zsa* '100 kilograms'.

7. Isolated cases. Three morphemes have root variants pronounced with a geminate consonant, even though they are written with a single grapheme,[20] *egy* 'one', *kisebb* 'smaller', and *lesz* 'he, she, it becomes'.

Symbolized:
egy → eggy
lesz → lessz
kisebb → kissebb

B. PHONOLOGICAL DEVIATIONS

8. *ly* → j. As mentioned above, *ly*[21] is identical in pronunciation with *j*, i.e. here we have a two-to-one relationship between grapheme and sound. It would be possible to dispense with this

17. The case of *ronts* 'destroy' is the only outright case of misinterpretation in Hungarian orthography; *ronts* should be written *roncs* (cf. the recent nominal formation, written *roncs* 'wreck').
18. *Utca* is sometimes spelled *ucca*; here, *út* 'road' has intruded from folk etymology.
19. Historically, the Old Hungarian [x] is retained only before vowels, as the laryngal [h]; in recent loanwords, however, the [x] has reappeared in root (and syllable) final position, e.g. *sah* 'shah', *potroh* 'abdomen (of an insect)', *pech* 'bad luck' [pɛhh]. Sometimes the [h] is optional, e.g. *düh* 'anger' [dü] or [düh]. There is no problem for the purposes of this paper, since it is only underlined *h* that is not pronounced.
20. When followed by a consonant, the *gy* of *egy* is short, in accordance with Rule 1, e.g. *egyben* 'by one'; also in a number of derivatives: *egyed* 'individual', *egyedül* 'alone', *egyetem* 'university' (in these cases, of course, no underlining is required.) In the positive form, the *s* [š] of *kis* 'little' is short. Likewise, in the other persons, the *sz* [s] of *lesz-* is short, e.g. *leszek* 'I become'.
21. The grapheme *ly* continues an older palatal lateral [lʲ]; its reflexes in dialects vary, e.g. *király* 'king' [kirāj] (the standard pronunciation), [kirāl], and [kirālʲ]. Therefore one might argue that in the native *Sprachgefühl* it did not coincide with *j* and consequently the difference in spelling was retained in spite of those who advocated replacing the *ly* by *j*.

rule, but for the sake of neatness it is included. Note that the distinction sometimes serves to differentiate words, e.g. *fojt* 'he strangles', *folyt* 'it was flowing'. They may also be combined e.g. *folyjon* 'let it flow'.

Symbolized: *ly* → j
lly → jj

9. Nasal adjustments of *n*[22] The grapheme *n* represents the general „all-purpose" nasal [n] and is adjusted to the following obstruent, e.g. *kinban* → kimban 'in severe pain', *szenved* → szemved 'he suffers', *konty* → konyty 'chignon', *rongy* → ronygy 'rag'.

Symbolized: *n* → m before *p, b, f, v*
n → ny before *ty, gy*

10. The rarely occurring voiced affricates *dz* and *dzs* are always long when permissible according to the distributional rules of Hungarian, i.e. intervocalically and in word-final position,[23] e.g. *edz* → eddz 'he trains', *hodzsa* → hoddzsa '(Turkish) wise man' (but *lándzsa* 'lance').

Symbolized: *dz* → ddz intervocalically and in word-final position
dzs → ddzs intervocalically and in word-final position

11. There are two phonemes in Hungarian which have no graphic representation:[24] short unrounded open [a] as in *Svájc* 'Switzerland', and a labial semivowel [w] as in *autó* 'car'. These cases, of course, cannot be indicated by means of the Hungarian orthography; however, attention will be called to them by reference to examples.

Symbolized: *á* → as the *á* in *Svájc* 'Switzerland'
u → as the *u* in *autó* 'car'.

Here again, → symbolizes re-interpretation, rather than respelling; cf. Rule 6.

C. IDIOSYNCRASIES

12. Deviations in the standard orthography occur in archaic native forms, especially in family names, and sometimes in place names. The most common cases are *cz* → c and, at the end of names, *y* → i.

22. Nasals pose an interesting problem in Hungarian and have attracted wide attention. (Cf. Eli Fischer-Jørgensen in *Actes du huitième congrès international des linguistes*, Oslo, 1958, p. 475.) The various structural approaches since Trubetzkoy have great difficulties even in formulating the problem. It is easy, of course, to give mechanical rules for the generation of these sounds. I think the solution lies in the fact that the palatal articulation differs essentially from the other buccal articulations. (Cf. my „X-Ray Sound Motion Picture Film and Some Phonological Questions of Hungarian", *Ural-Altaische Jahrbücher*, vol. 36, 165, pp. 31–38.)
23. Since *dz* and *dzs* are always pronounced as geminates when distributionally possible, one might regard them as unmarked, and derive the single affricate from Rule 1. I have not followed this treatment, though much can be said in favor of it, because I wanted to stick closely to the „substance", e.g. [ɛʒʒük] can be written in several ways: *edzük* 'we train', *eddzük* 'let us train it'(accepted since 1954 as the prescribed spelling, interpreting the [dz] as an affricate; earlier it was spelled *edzzük*, treating it as a verb ending in *z*). There is one example of final *dzs*, *bridzs* 'bridge (game)'. (In some pronunciations the derivative suffix *-ódz/-ődz* is pronounced with short [ʒ].)
24. Cf. note 9, above.

Symbolized: cz → c
y → i

13. Foreign elements. In foreign words the letters, *q*, *w*, *x*, independent *y*, and the digraph *ch* occur. Sometimes the spelling of the original language is retained, e.g. *Marx, hexameter, Darwin, technika. ch* is always short in consonant clusters and in initial position; intervocalically and in word-final position the digraph *ch* is sometimes short, sometimes geminate; the few cases of gemination have to be individually treated (cf. 15. below). E.g. *almanach* [ɔlmɔnɔh] 'almanac', but *pech* [pehh] 'bad luck (slang)'.

Symbolized: x → ksz
qu → kv
w → v
ch → h

14. There are a number of provincial colloquial, and fast pronunciations. In standard pronunciation, the more common cases are: *lj* → jj. e.g. *éljen* → éjjen 'may he live!'; *zs, szs* → ss, e.g. *község* → kösség 'village', *egészség* → egésség 'health'; $C_1C_1 \rightarrow C_1$, e.g. *kommunizmus* → komunizmus 'Communism', *milliő* → milió 'million'.

The cases of colloquial, substandard, or fast-contracted pronunciations are of no pedagogical significance in our context.[25]

15. The most radical way of dealing with idiosyncracies is by complete respelling, e.g. *Thewrewk* → Török, *Shakespeare* → Sekszpir, *Washington* → Vasinkton.

The conversion rules, which are summarized in the accompanying chart, allow any Hungarian text to be respelled or re-interpreted within the system of the native orthography itself, thus

25. In actual usage, there are standard pronunciations which differ from their orthographical representation, e.g. variations in consonant quantity and the quantity of the high vowels [ī], [ū], [ű]. Also, certain compounds or less frequent technical terms are usually not pronounced in accordance with the rules indicated above, e.g. *kétség* [kēččēg] 'doubt' vs. *rétség* 'meadowland' (a technical geographical term), usually pronounced [rētšēg]; *hatszor* 'six times' vs. *szétszór* 'scatter around' (the first is always pronounced [hɔccor], the second usually [sētsɔ̄r]). Likewise [ts] in *Margitsziget* 'Margaret Island'. For the cases where no change takes place, the solution is simple: no underlining is needed.

In addition to these standard variations there are also sub-standard cases, e.g. gemination: *erössen* for *erősen* 'strongly'; or contractions: *naccságos* for *nagyságos* 'honorable' (term of address); *mit csinál?* 'what is he doing? ' as [mičināl]. These are not part of the standard language and do not have to be indicated, but they can be handled by complete respelling; cf. Rule 15.

It is interesting to note that bookish pronunciation, which disregards many of the changes described in the rules above, occurs quite often. (I collected a dialect term from Southern Hungary for this: *irásilag beszél* 'he speaks according to the writing.')

26. The phonological changes underlying the conversion rules in script are as follows:

Qualitative Changes
Laryngal
Loss of [h]–Rule 5
Voicing switch–Rule 2
Apical
Adpalatalization–Rule 3
Adaffrication–Rule 4
Nasal
Palatal adjustment–Rule 9
Labiodental adjustment–Rule 9
Quantitative Changes
Degemination–Rules 1, 3, 4, 7, 10, (13), (14)

For explanation and examples, consult the rules cited.

providing an exact pronunciation for any work in the text in accordance with the phonemic values assigned by the phonemic principle in C.1 above.[27]

In practice, the method works as follows: Graphic segments which do not follow the phonemic principle described in C.1 are marked. Only one kind of marking, that of underlining, is used for all segments which are to be reinterpreted by a conversion rule. A single marking convention is sufficient since the situations covered in the chart of conversion rules are disparate, mutually exclusive and non-overlapping. Therefore the applicable rule can be selected without raising the question of ordering. Underlining instructs the reader to seek out a rule in the chart which applies to it and to convert accordingly. For example, *háztól* 'from the house' would be changed according to Rule 2, z → sz, resulting in the phonemic-orthographic spelling *hásztól*.

Sometimes two operations have to be performed on certain letter(s) in the same segment. The application of two rules is indicated by double underlining of the appropriate segment, with the stipulation that the second conversion operate upon the result of the first; e.g. *szebbtől* 'from the more beautiful' becomes according to Rule 1,2 bb → b, and according to Rule 2, b → p, resulting in the phonemic-orthographic spelling *szeptől*. Sometimes the underlinings are of unequal length, e.g. *játssza* 'let him play it'. Here Rule 1 reduces the geminate ssz to sz, then Rule 4 converts the resulting tsz to cc, resulting in the spelling *jácca*.

To demonstrate the application of the conversion rules to a connected text there follow first a few examples, prepared specifically for this paper and presenting a concentration of instances, then a page from my *Hungarian Reader* (Stockholm, 1938, p. 47), the beginning of a short story by Géza Gárdonyi (1863–1922). The popular-narrative style of this story contains a higher percentage of conversion rule applications than an expository or poetic text and should serve to show how the application appears in practice.

The summary chart of the conversion rules faces the Sample Text. In a book it could be attached as a fold-out page, to be used with the entire textbook, without requiring that the user constantly turn to the location of the chart in the text.

EXAMPLES

Orthography:	*a kisebbség és többség közti egészséges viszony*
Rules:	7 1,2 1,2 2 6,14
Respelling:	a kissepség és töpség köszti egésséges viszony
	'the healthy relationship between minority and majority'
Orthography:	*méhek a méhkasban*
Rules:	5 2
Respelling:	méhek a mékazsban
	'bees in the beehive'
Orthography:	*kétszer látja*
Rules:	4 3
Respelling:	kéccer láttya
	'he sees it twice'

27. These rules are literally rewrite rules, to use Chomsky's term, except for Rules 6 and 15, which are re-interpretations rather than conversions. From the viewpoint of this analysis I see no reason to distinguish between the two types of rules.

The notation can also be used for statistical purposes, e.g. the statistical analysis of the Hungarian consonants, an area where the data are particularly unreliable because of unclarity concerning morpho-phonemes, sounds, letters and graphemes.

In the orthography the underlying morphophonemic shapes are written; they represent the language sign as a conveyer of meaning. The method employed here allows these signs to be converted and to emerge on the plane of expression.

Orthography:	*ne rontsd el!*	
Rules:	4,2	
Respelling:	ne rondzsd el!	
	'don't destroy it!'	
Orthography:	*folyjon*	
Rules:	8	
Respelling:	fojjon	
	'let it flow'	
Orthography:	*sokban*	*sokkban*
Rules:	2	1,2
Respelling:	sogban	sogban
	'in many'	'in shock'

SAMPLE TEXT

Numbers in the right-hand margin refer to the rules in the Conversion Chart

21. Festő a falun.

– Gárdonyi Géza. –

Ma, ahogy kilépek az utcára, látom, hogy a gyerekek mind az alvég felé futnak. Megállítom a Burucz-gyereket, hogy mi történt? 4 2 12

– Egy úr van a faluvégen, – mondja zihálva, kipirosodottan, – képet csinál. 7 3

Egy úr, aki képet csinál; – nem lehet egyéb, csak festő. Magam is arrafelé néztem. Csakhamar rá is akadtam. 7 2 2

Ott ült a libalegelőn, fűzfák között. Előtte egy kis háromlábú festőállvány. Mögötte Köncöl Pista hadonászott a botjával, hogy távol tartsa a gyerekeket. Azok persze felmásztak a fára; onnan lesték, hogyan készül a csuda? Ki hitte volna, hogy olyan apró pálcikákkal ízélik a képet. Egy hét mulva minden gyerek fest a faluban. 2 2 1 3 4 8 2

A festő fiatal, szőke ember volt, afféle hosszúhajú szúnyoglegény, aminőket mindenfelé lehet látni nyáron az országban. A fején széleskarimájú puha kalap; a kabátja meg bársonyból való, olaszos. 9 3

De magyar fiú volt. Amint figyelmeztették, hogy a tanító jön, letette az ecsetjét és felkelt. 2 3

– Réz István vagyok, – mondotta – Münchenből jöttem haza az őszre, egynéhány stúdiumot csinálok itthon. 13 9 1

– Hogy vetődött ide, a mi kis elrejtett falunkba? 2

– Magam sem tudom. Járok jobbra-balra, amerre a szél hord. Tanulgatok. Még egy heti időm van, aztán visszamegyek. 1 2

– Tessék folytatni, azért beszélgethetünk. 8

A rétnek egy darabját festette: a hidat, a nyárfát, meg egynéhány fűzfát, a víz mellett. Fölvette az ecsetjét és föl-fölpillantva festett tovább. 2 3

– Látszik, – mondotta – hogy itt még sohase járt magamféle ember. A falu tele van szebbnél-szebb tanulmányfejjel, de akárkinek szóltam eddig, hogy legyen modellem, ijedten tiltakozott. 4 1 2

– Isten tudja, mit értenek a modell szón. A maguk nyelvén kell azokkal beszélni, jó festő uram. 3

– Beszéltem én az ő nyelvükön is. Egy barnaképű fiúcska végre leküzdötte egy hatosért a félelmét, de azt is elvitték tőlem. 2 2

CHART OF CONVERSION RULES

1. $\underline{C_1}\,\underline{C_1} \rightarrow C_1$ preceded or followed by a consonant.

2. $\begin{matrix}\underline{p} & \underline{t} & \underline{k} & \underline{ty} & \underline{f} & \underline{sz} & \underline{s} & \underline{c} & \underline{cs} \\ \underline{b} & \underline{d} & \underline{g} & \underline{gy} & \underline{v} & \underline{z} & \underline{zs} & \underline{dz} & \underline{dzs}\end{matrix}$ before a grapheme in the switching set and *h*

3. $\left.\begin{matrix}\underline{t}\ \underline{ty} \\ \underline{d}\ \underline{gy} \\ \underline{n}\ \underline{ny}\end{matrix}\right\} + \underline{j} \rightarrow \begin{cases}\text{ty after a consonant, tty otherwise} \\ \text{gy after a consonant, ggy otherwise} \\ \text{ny after a consonant, nny otherwise}\end{cases}$

4. $\left.\begin{matrix}\underline{t}\ \underline{tt} \\ \underline{d}\ \underline{dt}\end{matrix}\right\} + \begin{cases}\underline{sz}\ \underline{z}\ \underline{c} \rightarrow \text{c after or before a consonant, cc otherwise} \\ \underline{s}\ \underline{zs} \rightarrow \text{cs after or before a consonant, ccs otherwise}\end{cases}$

5. $\underline{h} \rightarrow \phi$ in root-final position, except before a vowel

6. $\underline{C} \rightarrow$ the grapheme C is not part of a digraph or trigraph

7. *egy* → eggy
   lesz → lessz
   kisebb → kissebb

8. *ly* → j
   lly → jj

9. *n* → m before *p, b, f, v*
   n → ny before *ty, gy*

10. *dz* → ddz intervocalically and in word-final position
    dzs → ddzs intervocalically and in word-final position

11. *á* → as *á* in *Svájc* 'Switzerland'
    u → as *u* in *autó* 'car'

12. *cz* → c
    y → i

13. *x* → ksz
    qu → kv
    w → v
    ch → h

14. *lj* → jj
    zs, szs → ss
    $\underline{C_1}\,\underline{C_1} \rightarrow C_1$

15. complete respelling

THE IMPERATIVE IN HUNGARIAN
(SPOKEN AND WRITTEN)

O. PRELIMINARY REMARKS

0:1 This paper describes the Imperative in Hungarian, both A) Spoken, and B) Written, and states C) the correlation between the two.[1]

0:2 The Imperative is the most complex grammatical category in Hungarian for the following reasons:

a) The paradigm is suppletive (the only clear example in Hungarian morphology); the Definite Singular 2nd Person Short Form has *no* mode-marker (*zero*-morpheme), while all other forms have an *explicit* imperative morpheme (though sometimes fused with the final consonant of the preceding morpheme.)

b) There is a special correlation, for most verbs,[2] in the Singular 2nd Person – unique in Hungarian conjugation – between a Short Form, expressing „command," and a Long Form, expressing „precation"; this gives 15 „slots" in the imperative paradigm, instead of the usual 13.[3]

c) The morphophonemic alternations are more complex than in any other morphological category.[4]

d) There are special morphemic alternations in the morpheme preceding the imperative morpheme, which occur only in this category.

e) Also, there are a few verbs, defective only in the Imperative: such as: *šikl(ik)* 'glide,' *čukl(ik)* 'have hiccups,' etc.[5]

A. SPOKEN HUNGARIAN

1. *Theoretical Remarks*

1:1 The basic problem in the analysis of the Imperative is to account for the concatenation between the Imperative morpheme and the immediately preceding morpheme – or, rarely, morphemes –, specifically its last phoneme. This segment is either the root – there are some

1. The general meaning of the Imperative in Hungarian is „directed action": either command as in *mɛnj* 'go (thou)!' or goal as in (*kɛ̄rt hod*^{j}) *mɛnj*^{j} (he asked that) you shall go.' Therefore, it has a full paradigm.
2. The only exceptions are a few verbs which do not have a short form in the non-definite form, e.g. *ɛd'ē!* 'eat!,' but even most *-ik* verbs allow a differentiation, cf. *ɔ hud*^{j}*d*^{j} 'sleep!' vs. *ɔ hud'd'āl* '(please) sleep!,' though in standard literary and school use the latter is recommended.
3. The three person categories combined with the number categories, Singular and Plural, and a further differentiation between Definite (referring to a „definite object in the third person") and Non-Definite. In addition, there is an Inclusive form in the singular first form I → you,' e. g. *lātlɔk* 'I see you.'
4. The Instrumental and Translative have more alternants, but the rules of their phonemic structures and the principle of selection are much simpler.
5. Defectiveness, of course, is meaningful only in a total verbal paradigm. Nevertheless, defectiveness in an „agglutinative" language is a significant fact, since usually all morpheme combinations are permissible.

special radical alternations before the Imperative – or a derivational suffix. It will be called *premorpheme.*

1:2 The method to establish the morphs in a given phonemic chain utilizes paradigmatic comparisons of both morphemes. This step may lead to one of three results at this stage of the analysis:

1:21 Some phonemic material is left over between the two morphs (*connective*)[6]; e.g. *hāz/ɔ/m* 'my house,' cf. *hāz* 'house' and *-m* 'my.' The Imperative never has any connective.

1:22 There is a definite cut between the two morph-stretches (*juxtaposition*); e.g. *hɔjō/m* 'my ship,' cf. *hɔjō* 'ship,' and *-m* 'my,' or *vār/j* 'wait,' cf. *vār* 'wait' and *-j* '(explicit imperative mark).'

1:23 Some phonemic materials belong to both paradigms (*overlapping* or *fusion*); e.g. *ērvɪl* 'by the argument', cf. *ērv* 'argument' and *-vɪl* 'with,' *mondd* 'say it!' cf. *mond* 'say' and *-d* '(2nd Person Definite Short Form Imperative).'

1:3 The class of morphs, united as a morpheme, do not form a well-ordered set. The constituent morphs may differ, however, as to the generality or restriction of their occurrence, as to their frequency or scarcity, as to their predictability or unpredictability from the distributional environment, and as to their phonetic freedom or motivation. The alternant showing greatest generality and least restriction is called the *basic* (unmarked) alternant.[7]

1:4 The rule according to which the appropriate morph is chosen of the set of morphs, included in the morpheme, is called *selection.* Selection involves two factors: a) the *role of the participating morphemes* in the selection, and the *direction* of the selection, and b) the *principle* according to which the selection takes place.

1:41 The selection may be dependent solely on one of the morphemes (*single selection*), or both morphemes may participate in the determination of the appropriate morphs (*dual selection*). The direction of the selection may be *progressive, regressive,* or *reciprocal.*

1:42 The factors involved in the selection may be *phonological,* or *non-phonological* (usually called *morphemic*), or *combined.* The phonological selections may be *automatic,* or *non-automatic.* Among the automatic selections the *phonetically motivated* morph selections are of particular interest, where a phonetic long component is established in the phonemic chain, e.g., unvoicing in *hāstōl* 'from (the) house,' cf. *hāz* 'house,' and *-tōl* 'from.' If there are no phonemic categories which determine the selection, we have to resort to lists of morphemes (this is obviously less economical).

2. *Premorphemic Alternations*

2:1 General morphophonemic alternations in Hungarian, which occur in the Imperative as well, are the following:

2:11 All morphemes ending in obstruants (stops, fricatives, or affricates), in which voicing is distinctive, have a paired alternant differing in the opposite feature of voicing. In the case of the Imperative, this principle is applicable in the Short Form of the Second Person Singular Definite Form, when the basic alternant of the premorpheme ends in a voiceless obstruant.

2:12 All morphemes ending in a geminate shorten this into a single phoneme if the suffix begins in a consonant, e.g. *šɔktōl* 'from chess,' cf. *šɔkk* 'chess' and *-tōl* '(abl.).'[8]

2:2 Among the morphemic alternations occurring in the Imperative only, there are three types: those which occur in the entire Imperative paradigm; those which occur only in connection with the explicit morpheme; and those which occur only and solely before the Short Form of the Indicative 2nd Person Singular Indefinite.

6. In Hungarian, it is always possible on the basis of internal evidence to assign the leftover phoneme to either the preceding or following morpheme. As a principle, we would not set up a connective-morpheme.
7. The question of *unmarked* vs. *marked* is an unclarified notion in linguistics. The above is not meant to be a formal definition.
8. In some pronunciation this reduction is not complete.

2:21 Alternations occurring in the entire Imperative paradigm comprise the verbs ending in *-t* (*-t* is a very common morpheme, indicating 'activity').
2:211 Verbs ending in *-st* and *-ʃt* have an alternant without the final *-t* in the imperative, ending thus in *-s* and *-ʃ*.
2:212 Morphemes ending in *short vowel + -t*, and also the roots *lāt* 'see,' *bočāt* 'forgive,' and *lōt* (occurring only in the „twin word" *lōt-fut* 'run around aimlessly') have an alternant ending in *-ǰ*.[9]
2:213 Morphemes ending in *consonant + -t* or *long vowel + -t* (except the three roots just listed) have an alternant ending in *-č*.
2:214 Four verbs having an alternant ending in *-cc* have in the imperative an alternant ending in *-ʃ*, though the unchanged alternant is also heard: *tɛcc(ik)* 'please,' *lācc(ik)* 'seem,' *həllɔcc(ik)* 'be heard,' *mɛcc* 'cut,' (of course, the general morphophonemic rules of 2:1 are applicable).[10]
2:22 Only before the explicit imperative morpheme do the following alternations occur:
2:221 Morphemes having an alternant ending in the appicals *-d* and *-n*, also have an alternant ending in the corresponding palatal, in *-d'* and *-ń*. (Thus *-t* and *-d* do not behave identically in the Imperative.)
2:222 Morphemes having an alternant in *-l*, in *Colloquial* Hungarian, but not in the literary usage, also have an alternant without this *-l*, cf. *ēlɟɛn* 'vivat!' and the more common *ēljɛn*.
2:23 Three monosyllabic roots have an alternant ending in *-ē* before the Short Form of the Indicative 2nd Person Indefinite Singular: *tē-* 'do,' *vē-* 'take,' and *lē-* 'become'.

3. *Imperative Alternants*

3:1 The *zero-suffix* occurs in the short form of the Singular 2nd Person Definite.
3:2 The *explicit* morpheme has a large number of morphs, all articulated in the pre-lingual zone. These are: either prepalatals, produced by the coronal muscles of the tongue; or dental-alveolar fricatives or affricates, produced by the apical muscles of the tongue. The enclosed Chart I. sums up the phonemic shapes of the various morphs.

Chart I.
The alternants of the explicit imperative morpheme

	Apical-dentialveolar		Coronal-prepalatal		
	Fricative	Affricate	Occlusive		Fricative-Semivowel
			Nasal	Oral	
Hissing	-s/-z	-(c)/-(z)	-ń	-d'-d' d'	-j -jj
Hushing	-ʃ/-	-č/-			

/ indicates voicing correlation
- indicates that the correlative pair does not occur in the imperative
() indicates that the alternant occurs only in fusion.

9. The *-ā* before the *-t* is not a sufficient condition for this alternation; e.g. *tāt* 'open widely' is treated like the verbs ending in *long vowel + t*.
10. The ending *-cc* is clearly bimorphemic – except in the case of *tɛcc(ik)* 'it seems, it pleases' – and the orthography shows this. (Historically *tɛcc(ik)* is also a case of bimorphemicity but descriptively this is not clear. Its connection with the politeness expression *tɛssēk* 'please,' is not felt any more.) The occurrence of the *-cc* is not a sufficient condition; e.g. *jācc(ik)* 'play' does not show this alternation in the Imperative. In some usage also the four verbs listed remain unchanged except for morphophonemic adjustments of voicing and simplification.

4. *Morphophonemic Concatenation*

4:1 The *zero-morpheme* is added to the morpheme-alternant ending in a fricative or affricate, if there is such an alternant. If there are two alternants differing in voicing, the voiced alternant is selected. (Postconsonantally the person suffix is *-d,* postvocalically *-dd,* but this is not a problem concerning the Imperative. There is always a distinction between the imperative and the corresponding indicative form: e.g. *ɔdod* 'you give it,' *ɔdd* 'give it!,' *lātod* 'you see it,' and *lāžd* 'see it!.')

4:2 The selection of the appropriate premorphemic alternant and the appropriate explicit imperative morpheme is done mainly by phonemic selection, but sometimes the selection is morphemic. There are the following cases (tabulated in Chart II.):

Chart II.
Morphemic concatenation in the Imperative

Explicit Imperative Alternant	Cooccurring premorpheme
-j	p/b l j k/g; long vowel f/v r m
-jj	*jö-*; (l ~ ∅); (long vowel)
-d'	d'; d ~ d'; *tɛ-vɛ-lɛ-vi-ɛ-i-; tē-vē-lē-**
-d'd'	hi-
-ń	ń; n ~ ń
-s	s; st ~ s
-z	z
-š	š; št ~ š; t ~ š; cc ~ š; *lāš-, bočāš, lōš-*
-(c)	cc
-(ʒ)	ʒʒ
-č	t ~ č

Underlining indicates morphemic selection (the rest is phonemically determined). The phonemes refer to the final portion of the premorpheme.

**tē-vē-lē* occur with the short form, *tɛ-vɛ-lɛ* occur with all the other forms.

~ morphophonemic alternation of the final portion of the premorpheme with the basic alternant listed first

() occur only in fusion

/ voicing correlation

4:21 Morphemically determined are the following selections: *jö-* 'come,' selects *-jj; hi-* 'believe,' selects *-d'd'.* The other six monosyllabic verbal roots, ending in vowel, select *-d':* *tɛ-* 'do,' *lɛ-* 'become,' *vɛ-* 'take,' *vi-* 'carry,' *ɛ* 'eat,' and *i-* 'drink.'[11] The selection is progressive.

4.22 The Indicative Second Indefinite Short Form marker *-d'* selects the root ending in *-ē* in three cases: *tēd'* 'do,' *lēd'* 'be,' and *vēd'* 'take.' Here the selection is regressive.

4.23 Progressive selection takes place after the few verbs ending in long vowel: e.g. *nōj* 'grow' (in these cases in colloquial usage *-jj* may be heard).[12]

11. *ɔls(ik)* 'sleep,' *fɛks(ik)* 'lie down,' *ńuks(ik)* 'rest' can better be treated in the Imperative as a predictable alternation from their alternant stems ending in *-d.*

12. The phonemic status of *-j* and its role in the quantitative opposition is not clear in Hungarian phonemic analysis.

4.24 After labials (*p, b, f, v, m*), after *j*, after *r*, and after the dorsals (*k* and *g*), the selected morph is *-j*; also after *l* in formal speech. Here the selection is progressive.
4.25 A reciprocal dual selection takes place in all the other alternants; these agree in voicing, in the character of the stop (oral or nasal), and in the kind of friction (fricative or affricate). Here the premorpheme determines the selection among the imperative morpheme alternants, and the imperative morpheme selects the alternant corresponding in friction or palatality among the alternants of the premorpheme.[13]
4.26 The premorpheme ending in *-l* or *-ll* (this is a very common ending) the colloquial pronunciation drops the *-l* and usually *-ll* is added; e.g. *ēljɛ́n*, or *ēlljɛ́n* 'vivat! '
4.27 The basic alternant is *-j*.[14]
4:3 The phonemic result of this morphemic concatenation is either *juxtaposition* or *fusion*. Fusion occurs if the premorphemic alternant ends in a consonant cluster in which the second element is identical with the imperative alternant or with the person marker in the case of the *zero*-imperative, e.g. {bonč} + {č} → /bonč/. (*c*) and (ʒ) occur only in fusion.[15]

5. *Syncretism*

5:1 There are two types of syncretisms in the Imperative:
5.11 One is the result of morphemic fusion, as indicated above; e.g. *mond* 'say' and 'he says,' *vonz* 'attract' or 'he attracts,' *kɛ́zd* 'start' or 'he starts.'
5:12 The other type of fusion occurs between some forms of the indicative present definite paradigm, including the Plural First Person of the front group, and Singular Third Person and the entire plural in the velar group. The syncretisms include only the cases where there are no special Imperative alternants, e.g. *ǎšša* 'he digs it' and 'he shall dig,' *kērjük* 'we beg it' and 'we shall beg,' but *lāt't'ɔ* 'he sees it,' and *lāššɔ* 'he shall see it' are differentiated.

B. WRITTEN HUNGARIAN

6:1 In Written Hungarian the following graphic alternants of the Explicit Imperative morpheme occur: *-j, -jj, -gy, -ggy, -s, -sz, -z*, and *-dz*. The *zero*-morpheme poses no problem; the person marker, *-d*, or *-dd*, is added without a connective.
6:2 In the graphic segment preceding the Imperative morph, the following morphographemic alternations occur (cf. Chart III.):

13. The predictability is uni-directional; from the basic alternant *ɔd* and *hɔd'*, the Imperative is predictable, but not in reverse.
14. Basic is the form which shows the minimum degree of phonetic motivation.
15. The voiced hissing affricate occurs in two roots (*ɛʒʒ* 'train' and *pɛʒʒ* 'guess') and in the very common reflexive ending *-ōʒʒ* (alternating freely with *-ōzz*), and is always long, if distributionally possible.

Chart III.
The Explicit Imperative Morpheme in Writing

Graphemic Imperative Alternant	Cooccurring
-j	all cases, except the following:
-jj	*jö-*
-gy	*te-ve-le-vi-e-i-té-vé-lé*
-ggy	hi-
-s	s st ~ s t t ~ s tsz ~ s
-sz*	sz szt ~ sz
-z	z
-dz*	dz

Underlining indicates morphemic selection, the rest is graphemically determined. The letters refer to the final portion of the premorpheme.

~ morphographemic alternation; the basic alternant is listed first.

* sz + sz → ssz, dz + dz → ddz, except in hyphenation.

6:21 In the entire Imperative paradigm:

6.211 Verbal morphemes ending in *-st* and *-szt* have an alternant without the *-t*, thus ending in *-s* and *-sz*.

6.212 Verbal morphemes ending in *unaccented vowel-sign*, and the three roots *lát* 'see,' *bocsát* 'forgive,' and *lót* (in *lót-fut*) 'running around,' have an alternant in *-s*, *lás-*, *bocsás-*, and *lós-*.

6.213 Four verbs ending in *-tsz*: *látsz(ik)* 'seem,' *hallatsz(ik)* 'be heard,' *tetsz(ik)* 'please,' and *metsz* 'cut,' have an alternant ending in *-s*, *lás-*, *hallas-*, *tes-*, and *mes-*, though the unaltered form is also used by some.

6:22 Three roots, *té-* 'do,' *vé-* 'take,' and *lé-* 'become,' occur before the „short form" *gy*.

6:231 The selected premorphemic-alternant before the Imperative ends in *-s*, *-sz*, or *-z*, if choice is present.

6:232 The Explicit Imperative morpheme selected agrees with the final phone, *-s*, *-sz*, *-z*, and *-dz*.[16] Verbs ending in *-t* select *-s*. Isolated radical selections: *jö-jj* 'come,' and *hi-ggy* 'believe.'

6:3 Syncretism occurs between the Imperative and the corresponding Indicative Present form, e.g. *várja* 'he expects it, he shall expect it!.'

C. CORRELATION BETWEEN SPOKEN AND WRITTEN HUNGARIAN

7:1 The Imperative in Written Hungarian is much simpler than in Spoken Hungarian since a number of morphophonemic alternations which occur in speech, such as change in voicing or reduction in gemination, are not indicated in writing, because of the overriding morphophonemic (or morpheme-analytical) principle in Hungarian orthography. The only indicated changes are the quantitative distinctions, loss of phonemes, and alternations between stop and fricative.[17]

16. The graphemic rule about digraphs applies: *sz + sz → ssz*, and *dz + dz → ddz*, except in hyphenation, when both forms are written fully.

17. Thus there is a graphic distinction between the homophonous *roncs* 'wreck' and *ronts* 'destroy.' In my opinion, this is the only case in Hungarian orthography which clearly violates the morphophonemic principle; *rontsd* should be replaced by *roncsd* and *ronts* by *ronccs*, in analogy with spelling like *fuss* 'run.' This also applies to the spelling of verbs like *fűts* 'heat' which should be replaced by *fűccs*, and likewise *fűtsd* by *fűcsd*.

7:2 The script differentiates between two (d¹) in the Imperative according to their morphophonemic status; in the case of the morphonemically predictable (d¹) the script uses *-d-j* according to the morpheme-analytical principle, in the case of the morphemically selected (d¹) the alternant is written -gy in accordance with the phonemic principle in the orthography, cf. *ad-j* 'give!' and *té-gy* 'do!' (/ɒd¹ + d¹/ and / tẽ+d¹/).

7:3 Fusion is always resolved in script; cf. *vonz* 'he attracts' and *vonzz* 'attract!,' *mond* 'he says' and *mondd* 'say it!,' or *játszunk* 'we play' and *játsszunk* 'let us play,' *edzük* 'we train' and *eddzük* 'let us train.'[18]

18. There are also some sporadic formations of commanding nature, such as *ládd* 'see', *jer* 'come,' etc. *jövel* 'come' and a few interjectional formations, very often ending in the suffix *-sza/-sze*.

Addendum: The Imperative of the morpheme {mɑ-, mɑd¹- mẽd¹-, mɑn-, mɛn} 'go' is *mɑɲ-ɲ₂:* {vɒn, vɒd¹-, vɔt', vol-, vɔl-} 'be' is suppletive, {le-, lẽ-} 'become, (be)' being the suppletion.

BIBLIOGRAPHICAL NOTE

Sources of the papers in this publication are as follows:

1. „Script in the Framework of Grammatical Theory" is a revised version of a working paper submitted to the Burg Wartenstein Symposium, „Toward the Description of the Languages of the World" (August 1–8, 1970), published under the title „Comment: The Role of Script in Describing the Languages of the World" in *Visible Language* Vol. V, Number 1, Winter 1971, pp. 75–81.

2) „Select Bibliography on Script and Language" was compiled for a conference on reading problems sponsored by the National Institute of Child Care and Human Development in May 1971 at Belmont, Maryland and was published in *Visible Language,* Vol. VI, No. 1, 1972, pp. 79–80.

3) „Hungarian Script" is a chapter in my *Hungarian Reference Grammar,* submitted to the Office of Education of the United States in 1969. It is a reworking of a corresponding chapter in *Das ungarische Sprachsystem,* Stockholm, 1939.

4) „The Conversion of Script to Speech as Applied to Hungarian" is reprinted from the *Linguistic Reporter,* Vol. II, Number 5, October 1969, pp. 17–30 (= Supplement 23), with a few corrections and additions.

5) „The Imperative in Hungarian (Spoken and Written)" appeared in „American Studies in Uralic Linguistics", *Indiana University Publications, Uralic and Altaic Series,* Vol. 1, 1960. pp. 83–92.